OHIO RIVER BASIN PRECIPITATION FREQUENCY PROJECT

Update of *Technical Paper No. 40, NWS HYDRO-35* and *Technical Paper No. 49*

Fourteenth Progress Report
1 January 2003 through 31 March 2003

Hydrometeorological Design Studies Center
Hydrology Laboratory

Office of Hydrologic Development
U.S. National Weather Service
National Oceanic and Atmospheric Administration
Silver Spring, Maryland

April 2003

DISCLAIMER

The data and information presented in this report should be considered as preliminary and are provided only to demonstrate current progress on the various technical tasks associated with this project. Values presented herein are NOT intended for any other use beyond the scope of this progress report. Anyone using any data or information presented in this report for any purpose other than for what it was intended does so at their own risk.

TABLE OF CONTENTS

OHIO RIVER BASIN
PRECIPITATION FREQUENCY PROJECT

Update of *Technical Paper No. 40, NWS HYDRO-35* and *Technical Paper No. 49*

1. Introduction

The Hydrometeorological Design Studies Center (HDSC), Hydrology Laboratory, Office of Hydrologic Development, U.S. National Weather Service is updating its precipitation frequency estimates for the Ohio River Basin and surrounding states. Current precipitation frequency estimates for this area are contained in *Technical Paper No. 40* "Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years" (Hershfield 1961), *NWS HYDRO-35* "Five- to 60-minute precipitation frequency for the eastern and central United States" (Frederick et al 1977) and *Technical Paper No. 49* "Two- to ten-day precipitation for return periods of 2 to 100 years in the contiguous United States" (Miller et al 1964). The new project includes collecting data and performing quality control, compiling and formatting datasets for analyses, selecting applicable frequency distributions and fitting techniques, analyzing data, mapping and preparing reports and other documentation.

The project will determine annual all-season precipitation frequencies for durations from 5 minutes to 60 days, for return periods from 2 to 1000 years. The project will review and process all appropriate rainfall data for the project area and use accepted statistical methods. The project results will be published as a Volume of NOAA Atlas 14 on the Internet with the additional ability to download digital files.

The project will produce estimates for 13 states. Parts of nine additional bordering states are included to ensure continuity across state borders. The Susquehanna River and Delaware River Basins are included in the project area. The core and border areas, as well as tentative regions now used in the analysis, are shown in Figure 1.

Figure 1. Updated Ohio River Basin project area and new regional groups.
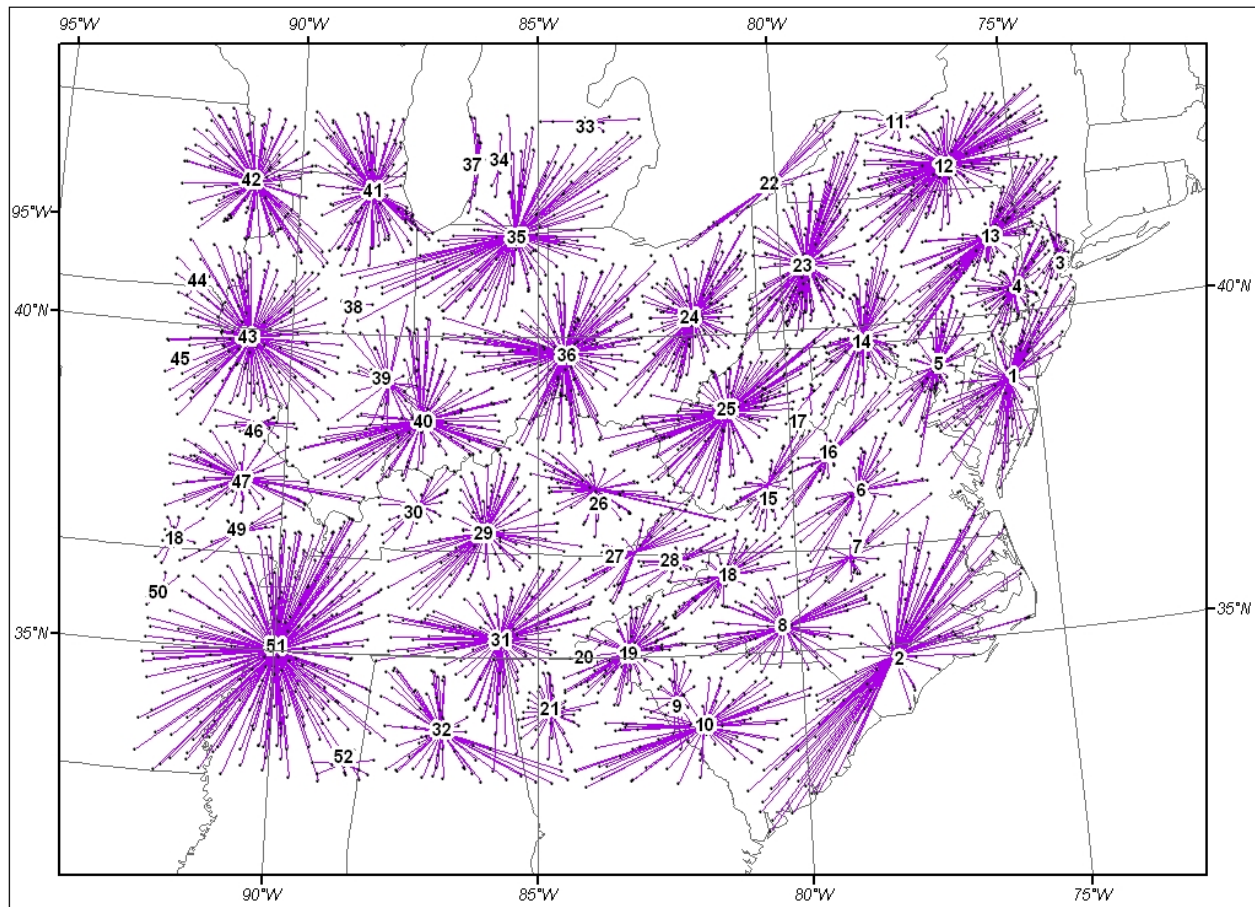
Ohio River Basin & Surrounding States Precipitation Frequency Regions

2. Highlights

Software to compute multi-day duration return frequencies was refined to accommodate different numbers of stations and to insure internal consistency between 24-hour and 48-hour durations. Software to calculate confidence limits for all durations and frequencies was refined. Additional information is provided in Section 4.1, Software Updates.

The original 16 regions have been subdivided into 52 smaller homogeneous regions. Given the large amount of stations and data incorporated into these regions, we are carefully examining the results, paying particular attention to maximum observed 1-day events that exceed 1000-year estimates. Return frequencies for shorter durations and multi-day durations were also computed for the 52 regions and examined for heterogeneity. Additional information is provided in Section 4.2, L-moment Analysis.

Generalized Extreme Value (GEV) is the best-fitting distribution in 28 of the 52 regions. However, the sensitivity of estimates to different distributions and the spatial continuity of best-fitting distributions is being examined. Additional information is provided in Section 4.3, Distribution Selection.

On February 19th, 2003, the Spatial Climate Analysis Service (SCAS) at Oregon State University delivered the draft 60-minute and 24-hour mean annual maxima (a.k.a. "Index flood") grids for the Ohio River Basin and surrounding States to HDSC. We are currently evaluating the 100-year 24-hour point estimates and grids before converting them to cartographic-quality maps for the peer review. Additional information is provided in Section 4.4, Spatial Interpolation.

The Precipitation Frequency Data Server (PFDS) is now operating at much faster speeds and is nearly ready for public release with the Semiarid Southwestern United States Precipitation Frequency Project. Additional information is provided in Section 4.5, Precipitation Frequency Data Server.

Progress towards the development of depth-area-duration (D-A-D) reduction relationships for areas from 10 to 400 square miles continues. An additional area in Hawaii has been added to the project. Additional information is provided in Section 4.6, Spatial Relations (Depth Area Duration Project).

HDSC presented four papers/posters at the 83rd American Meteorological Society Annual Meeting in February of 2003. Preliminary Ohio 100-year 60-minute and 100-year 24-hour point estimates at observing stations and a comparison between current 100-year 24-hour point estimates with TP40 were presented and well-received. Additional information is provided in Section 5.1, AMS Annual Meeting.

3. Status

3.1 Project Task List

The following checklist shows the components of each task and an estimate of the percent completed per task. Past status reports should also be referenced for additional information.

Ohio River Basin project checklist [estimated percent complete]:

Data Collection, Formatting and Quality Control [100%]:
- Multi-day
- Daily
- Hourly
- 15-minute
- N-minute

As data issues may arise in subsequent tasks, quality control is essentially a continuous process.

L-Moment Analysis/Frequency Distribution for 5 min. to 60 days and 2 to 1000 years [80%]:
- Multi-daily
- Daily
- Hourly
- 15-minute
- N-minute

Once the 52 regions are finalized, all durations and return frequency quantile estimates will be computed and adjusted for internal consistency.

Spatial Interpolation [25%]
- Create mean annual maximum (a.k.a. "index flood") grids with PRISM for all durations from 60-minute to 60-days
- Apply a precipitation frequency map derivation procedure, known as the cascade residual add-back (CRAB) procedure to create a total of 162 grids. The same procedure will be used to create 162 upper and 162 lower bound precipitation frequency grids (see Section 4.3, Spatial Interpolation).
- Apply project-wide conversion factor to the 1-hour precipitation frequency grids to calculate the n-minute (5-, 10-, 15-, and 30-minute) grids.

Draft grids of the 60-minute and 24-hour mean annual maximums have been delivered to HDSC.

Peer Reviews [0%]:
- External peer review of point precipitation frequency estimates
- External peer review of spatially interpolated grids

Data Trend Analysis [0%]
- Analyze linear trends in annual maxima and variance over time
- Analyze shift in means of annual maxima between two time periods (i.e., test the equality of 2 population distribution means)

Temporal Distributions of Extreme Rainfall [75%]
- Assemble hourly data by quartile of greatest precipitation amount and convert to cumulative rainfall amounts for each region
- Sort, average and plot time distributions of hourly maximum events by storm area, quartile and duration

Temporal distributions for 12- and 24-hour durations are complete. Work remains to compute the 4-day and 6-hour durations.

Deliverables [50%]
- Prepare data for web delivery
- Prepare documentation for web delivery
- Write hard copy of Final Report
- Publish hard copy of Final Report

Minor modifications have been made to the Precipitation Data Frequency Server (PFDS) to accommodate the Ohio Project.

Spatial Relations (Depth Area Duration Project) [67%]
- Obtain hourly data from dense-area reporting networks
- QC and format data from dense networks
- Compute maximum and average annual areal depth for each duration from stations for each network
- Compute maximum to average depth ratio for all durations and networks and plot
- Prepare curves of best fit (depth area curves) for each duration and network
- Combine all stations from all project areas to compute the ratio of maximum to average depth for all durations and networks and plot
- Examine differences in individual D-A-D curve plots for durations and different areas compared to those for combined area data plots

The D-A-D project is 2/3 completed. All areas to be used and tested in the D-A-D curve development have been selected and quality controlled. Software development to process the data and ultimately generate the D-A-D curves is 60% completed.

4. Progress in this Reporting Period

4.1 Software Updates

Minor inconsistencies were found in the Semiarid Project's daily and multi-day results that related to our automated processes. The inconsistencies typically occurred between 24-hour and 48-hour durations (i.e., the 24-hour estimate was greater than the 48-hour estimate). In all projects, 24-hour estimates are calculated from a combined dataset of hourly stations and daily stations. Multi-day durations are calculated from daily stations only. Therefore, it was necessary to modify the multi-day internal consistency software to accommodate the different number of stations between 24-hour and 48-hour.   This modification eliminated all observed inconsistencies.

Software to calculate confidence limits for all durations and frequencies was refined. An internal consistency check was built in to the software.

4.2 L-moment Analysis

The original 16 regions have been subdivided into 52 smaller homogeneous regions (see Figure 1 and Table 1) and the precipitation frequency estimates were computed. Given the large number of stations and data incorporated into these new regions, we are carefully examining the results, paying particular attention to a real data check for maximum observed 1-day events that exceed the 1000-year estimate at a station. There are 95 stations in 32 regions that have observed maximums greater than their 1000-year 24-hour estimate. Given the large number of stations and years of record in the dataset, it is statistically expected for some events to exceed the 1000-year estimate. Regions where maximum observed 1-day events exceed 1000-year estimates are being scrutinized and any additional subdivisions are made on a case by case basis. During the subdivision process, stations in regions may be re-grouped into smaller homogeneous regions based on climatology, topography, real data checks and statistical results.

Table 1. Number of stations in new regions.

new region	# daily	# hourly	total					
1	74	12	86		29	73	27	100
2	93	24	117		30	24	9	33
3	31	12	43		31	97	22	119
4	63	16	79		32	60	13	73
5	57	20	77		33	6	1	7
6	34	11	45		34	4	2	6
7	25	6	31		35	104	50	154
8	65	16	81		36	144	67	211
9	18	5	23		37	10	5	15
10	67	16	83		38	6	2	8
11	12	3	15		39	31	13	44
12	144	51	195		40	78	36	114
13	95	41	136		41	85	30	115
14	64	31	95		42	83	34	117
15	32	11	43		43	116	39	155
16	30	8	38		44	5	0	5
17	7	3	10		45	3	1	4
18	49	11	60		46	11	3	14
19	69	25	94		47	46	24	70
20	4	1	5		48	10	4	14
21	37	13	50		49	12	5	17
22	17	7	24		50	5	1	6
23	124	50	174		51	209	68	277
24	103	48	151		52	10	2	12
25	130	41	171		at-sites	10	1	13
26	49	15	64		totals	2788	973	3761
27	22	7	29					
28	31	10	41					

Note: Daily stations must have at least 30 years of data.
Note: Hourly stations must have at least 20 years of data.

Precipitation frequencies for shorter durations and multi-day durations were also computed for the 52 regions and examined for heterogeneity. During the multi-day analysis, 16 regions were identified where there are a different number of stations at different durations, not only at 24-hour and 48-hour. This occurs when the number of years of data at a station drops below the acceptable 30 years for a given duration. Possible reasons for this drop include:

1. The station has hand-entered monthly maximums resulting in 24-hour values only. Longer durations cannot be derived.
2. At a given station, a year does not have enough data to reliably extract an annual maximum for a given duration, but an observed value of that duration (usually 24-hour or 48-hour) exceeds a certain threshold and is therefore considered the maximum for that year.
3. There is not enough data to extract the longer durations such as 30-day and longer but there was enough data for shorter durations.

As mentioned, the existing internal consistency software is unable to accommodate different numbers of stations between durations. Therefore, to retain as much data as possible, the software will be modified to allow different numbers of stations at all durations. Each station where a discrepancy occurs was flagged and checked for data quality.

4.3 Distribution Selection

The best-fitting distribution for each region has been examined (Table 2). The best-fitting distribution is determined by three tests: Xtest, Root-Mean-Square-Error (RMSE) , and real data fit. The Xtest compares the average L-kurtosis and L-skewness of a region with theoretical distributions. The RMSE test measures the root-mean-square-error of real data L-kurtosis and theoretical distributions. The real data fit test compares the empirical frequencies to the probabilistic quantiles for the theoretical distributions (Lin and Vogel 1993). The Generalized Extreme Value (GEV) distribution is the best fit for 28 of the 52 regions; the Generalized Normal (GNO) distribution is the best fit for 13 regions; and the Generalized Log-Normal (GLO) distribution is the best fit for 8 regions. The sensitivity of estimates to different distributions and the spatial continuity of best-fitting distributions is being examined.

Table 2. Best-fitting distributions for each region.

Region (# of stns daily + hourly)	Best Fit		Region	Best Fit
Region 1 (74+12)	GEV		Region 27 (22+7)	GEV
Region 2 (93+24)	GEV		Region 28 (31+10)	GLO
Region 3 (31+12)	GEV		Region 29 (73+27)	GEV
Region 4 (63+16)	GNO		Region 30 (24+9)	GNO
Region 5 (57+20)	GNO		Region 31 (97+22)	GNO
Region 6 (34+11)	GNO		Region 32 (60+13)	GNO
Region 7 (25+6)	GNO		Region 33 (6+1)	GLO
Region 8 (65+16)	GNO		Region 34 (4+2)	GLO
Region 9 (18+5)	GLO		Region 35 (104+50)	GEV
Region 10 (67+16)	GEV		Region 36 (144+67)	GEV
Region 11 (12+3)	GEV		Region 37 (10+5)	GEV
Region 12 (144+51)	GEV		Region 38 (6+2)	GLO
Region 13 (95+41)	GEV		Region 39 (31+13)	GEV
Region 14 (64+31)	GEV		Region 40 (78+36)	GEV
Region 15 (32+11)	GEV		Region 41 (85+30)	GEV
Region 16 (30+8)	GNO		Region 42 (83+34)	GEV
Region 17 (7+3)	GLO		Region 43 (116+39)	GNO
Region 18 (49+11)	GNO		Region 44 (5+0)	GEV
Region 19 (69+25)	GNO		Region 45 (3+1)	GNO
Region 20 (4+1)	GEV		Region 46 (11+3)	GLO
Region 21 (37+13)	PE3		Region 47 (46+24)	GEV
Region 22 (17+7)	GEV		Region 48 (10+4)	GNO
Region 23 (124+50)	GEV		Region 49 (12+5)	GLO
Region 24 (103+48)	GEV		Region 50 (5+1)	GLO
Region 25 (130+41)	GEV		Region 51 (209+68)	GEV
Region 26 (49+15)	GEV		Region 52 (10+2)	GEV

4.4 Spatial Interpolation

On February 19th, 2003, the Spatial Climate Analysis Service (SCAS) at Oregon State University delivered the draft 60-minute and 24-hour mean annual maxima (a.k.a. "Index flood") grids for the Ohio River Basin and surrounding States to HDSC. The grids were subsequently converted into cartographic PDF maps that will be used during the upcoming peer review. The "index flood" grids allowed, for the first time, the derivation of draft 100-year 60-minute and 100-year 24-hour grids using the Cascade, Residual Add-back (CRAB) procedure. We are currently evaluating the 100-year 24-hour point estimates and grids before converting them to cartographic-quality maps for the peer review.

4.5 Precipitation Frequency Data Server (PFDS)

Last quarter the available disk space for the PFDS computer server was increased to 30 gigabytes, but the access speed to the expanded disk was hindered by a slow network connection. This quarter the slow connection was resolved by moving the PFDS computer server to a faster internal network. The PFDS is now operating at much faster speeds.

In preparation for the upcoming peer review, preliminary state-specific web pages were developed, which include station maps with color elevation as the background. The web pages also include several options for extracting, displaying and selecting point data. Although the PFDS was not populated with all of the draft point precipitation frequency estimates, some was added to allow for testing.

4.6 Spatial Relations (Depth-Area-Duration Project)

Progress continues in the development of geographically-fixed depth-area-duration (D-A-D) reduction relationships for area sizes of 10 to 400 square miles. The second phase of the programming to relate spatial relationships in precipitation data used in the development of the D-A-D curves is nearly complete and will be tested in April 2003 on two areas.

The purpose of the programming is to generate statistics that measure the variability in the annual maximum for a given duration in a given area or basin. Means and standard deviations among groups of five stations, normalized to the distance between stations, are computed. Mean areal depths (using annual maximum precipitation) in a basin are generated, in order to ultimately compute the ratios of mean annual maximum amounts at stations to the mean areal annual maximum. To quality check the software, data from an earlier study (NOAA Technical Report NWS 24, using this D-A-D development approach) is being duplicated to verify that the same statistics are being generated using our software development. This is the last major software development for the project.

An additional area in Hawaii has been added to the project. A total of 13 different geographic areas throughout the United States have been quality controlled and will be used in the project. The set of curves developed for each area will be tested for differences to determine if a single set of D-A-D curves is applicable to the entire U.S. Otherwise, separate curves for different regions of the country will be developed.

5. Issues

5.1 AMS Annual Meeting

HDSC presented four papers/posters at the 83rd American Meteorological Society Annual Meeting in February of 2003. In *Updating NOAA/NWS Rainfall Frequency Atlases*, we provided an overview of our approach; in *Updated Precipitation Frequencies for the Semiarid Southwest United States*, we presented preliminary 100-year 60-minute and 100-year 24-hour spatially interpolated maps and a comparison between current 100-year 24-hour point estimates and NOAA Atlas 2; in *Updated Precipitation Frequencies for the Ohio River Basin and Surrounding States*, we presented preliminary 100-year 60-minute and 100-year 24-hour point estimates from the Ohio project and a comparison between current 100-year 24-hour point estimates and Technical Paper 40; and in *NOAA/NWS Precipitation Frequency Data Server*, we presented the PFDS in detail. The papers were well received and the posters generated significant interest and anticipation of final publication.

6. Projected Schedule

The following list provides a tentative schedule with completion dates. Brief descriptions of tasks being worked on next quarter are also included in this section.

    Data Collection and Quality Control [Complete]
    Temporal Distributions of Extreme Rainfall [May 2003]
    L-Moment Analysis/Frequency Distribution [May 2003]
    Peer Review of Spatially Distributed Point Estimates [May 2003]
    Spatial Interpolation [June 2003]
    Precipitation Frequency Maps [July 2003]
    Web Publication [August 2003]
    Spatial Relations (Depth Area Duration Studies) [May 2003]

6.1 L-Moment Analysis

Regions will be finalized during the next quarter. This is expected to take roughly 4 weeks. Then a comprehensive L-moment statistical analysis will be completed for all datasets and durations.

6.2 Spatial Interpolation

During the next quarter, HDSC will derive the 100-year 60-minute and 100-year 24-hour precipitation frequency maps for the peer review. After addressing reviewer comments, HDSC will produce and send final mean annual maxima for all durations to be interpolated by PRISM.

6.3 Peer Review

A single peer review of the precipitation frequency point estimates and draft maps will occur during the next quarter. The review will include the point precipitation frequency estimates and associated confidence intervals for all durations (5-minute to 60-day) and all return frequencies (2-year to 1000-year). The review will cover all stations, even those outside the core area of the project. The purpose for including the non-core area is to provide continuous data across the exterior project area border. Comments pertaining to data in non-core areas will be addressed according to their influence to the core project area . The review will also include the spatially interpolated grids for the following:
1.   1-hour mean annual maximum maps ("index flood" maps)
2.   1-hour 100-year precipitation frequency maps
3.   24-hour mean annual maximum maps ("index flood" maps)
4.   24-hour 100-year precipitation frequency maps

6.4 Temporal Distributions of Extreme Rainfall

Temporal distributions will be completed during the next quarter.

6.5 Spatial Relations (Depth-Area-Duration Project)

Software for the D-A-D computations will be completed in the next quarter and the computations will be performed for 13 areas, and the resulting curves will be tested for differences to determine if a single set of D-A-D curves is applicable to the entire U.S. or whether curves vary by region.

References

Frederick, R.H., V.A. Myers and E.P. Auciello, 1977: Five- to 60-minute precipitation frequency for the eastern and central United States, NOAA Technical Memo. NWS HYDRO-35, Silver Spring, MD, 36 pp.

Hershfield, D.M., 1961: Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years, *Weather Bureau Technical Paper No. 40*, U.S. Weather Bureau. Washington, D.C., 115 pp.

Hosking, J.R.M. and J.R. Wallis, 1997: *Regional frequency analysis, an approach based on L-moments*, Cambridge University Press, 224 pp.

Huff, F. A., 1990: Time Distributions of Heavy Rainstorms in Illinois. Illinois State Water Survey, Champaign, 173, 17pp.

Lin, B. and L.T. Julian, 2001: Trend and shift statistics on annual maximum precipitation in the Ohio River Basin over the last century. Symposium on Precipitation Extremes: Prediction, Impacts, and Responses, 81st AMS annual meeting. Albuquerque, New Mexico.

Lin, B. and J. Vogel, 1993: A comparison of L-moments with method of moments, *Engineering Hydrology Symposium Proceedings*, ASCE, San Francisco, CA, July 25-30, 1993.

Miller, J.F., 1964: Two- to ten-day precipitation for return periods of 2 to 100 years in the contiguous United States, *Technical Paper No. 49*, U.S. Weather Bureau and U.S. Department of Agriculture, 29 pp.

Miller, J.F., R.H. Frederick and R.J. Tracy, 1973: Precipitation-frequency atlas of the western United States, *NOAA Atlas 2*, 11 vols., National Weather Service, Silver Spring, MD.

Myers, V.A. and R.M. Zeher, 1980: A Methodology for Point-to-Area Rainfall Frequency Ratios, NOAA Technical Report NWS 24, Office of Hydrology, National Weather Service, Silver Spring, MD.